

Research on **E**xemplary, **E**ffective, and **E**quitable Curriculum Materials



New York ECLIPSE

Enhancing Collaborative Leadership
for Improved Performance in
Science Education

Science Education Curriculum Showcase

Carol O'Donnell
Research Scientist

The George Washington University
Washington, DC

March 21, 2007

Outline

- Students' alternative conceptions
- Facts about science achievement
- Characteristics of exemplary science curriculum materials
- Results from research on the effectiveness of exemplary science curriculum materials:
 - Content analyses
 - Pre-experiments
 - Quasi- and true experiments
- Conclusions

Student Quotes from Science Tests, Essays, and Papers

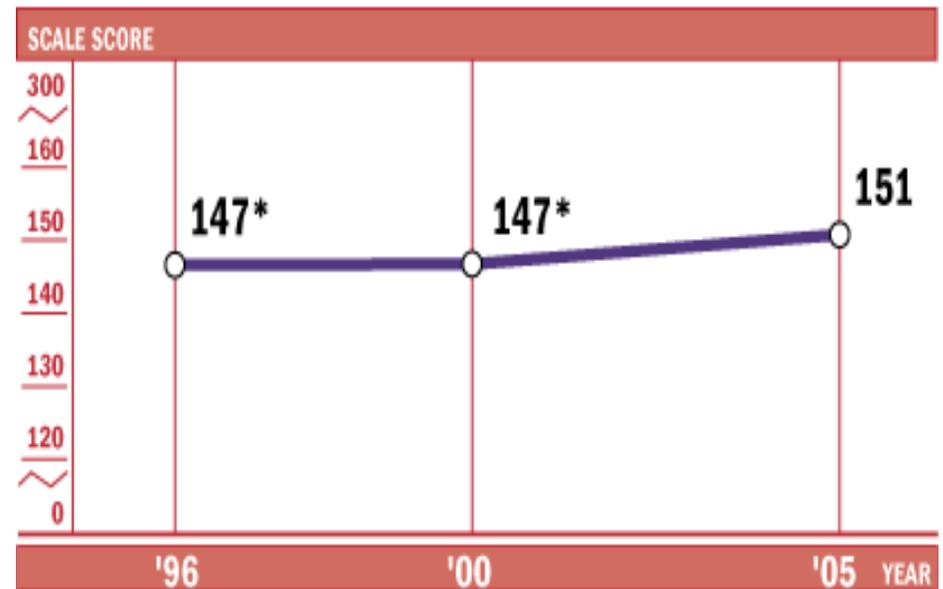
- "H₂O is hot water, and CO₂ is cold water"
- "When you breath, you inspire. When you do not breath, you expire."
- "Water is composed of two gins, Oxygin and Hydrogin. Oxygin is pure gin. Hydrogin is gin and water."
- "The pistol of a flower is its only protection against insects."
- "Germinate: To become a naturalized German."

Facts About Science Achievement

National Assessment of Educational Progress (NAEP)

- Average NAEP science scores for 4th graders were higher in 2005 than they were in 1996 or 2000.
- However, scores remained unchanged at 8th grade, and are significantly lower at 12th grade compared to 1996.

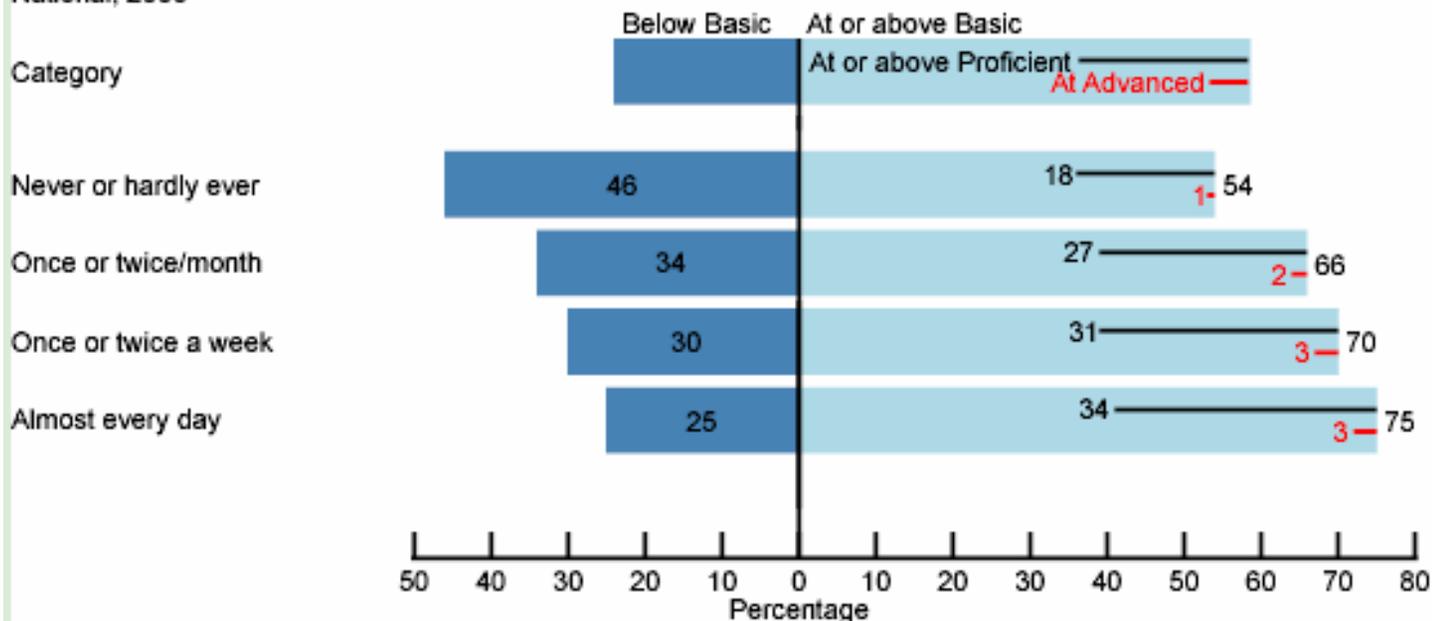
Percentage of 4th graders proficient in science



Sources: http://nationsreportcard.gov/science_2005/s0102.asp

2005 NAEP: National Achievement Levels for 4th Graders Who Do Hands-on Science

Percentages of students at or above each achievement level for science, grade 4
 Students do hands-on science activities (2005) [T090207]
 National, 2005



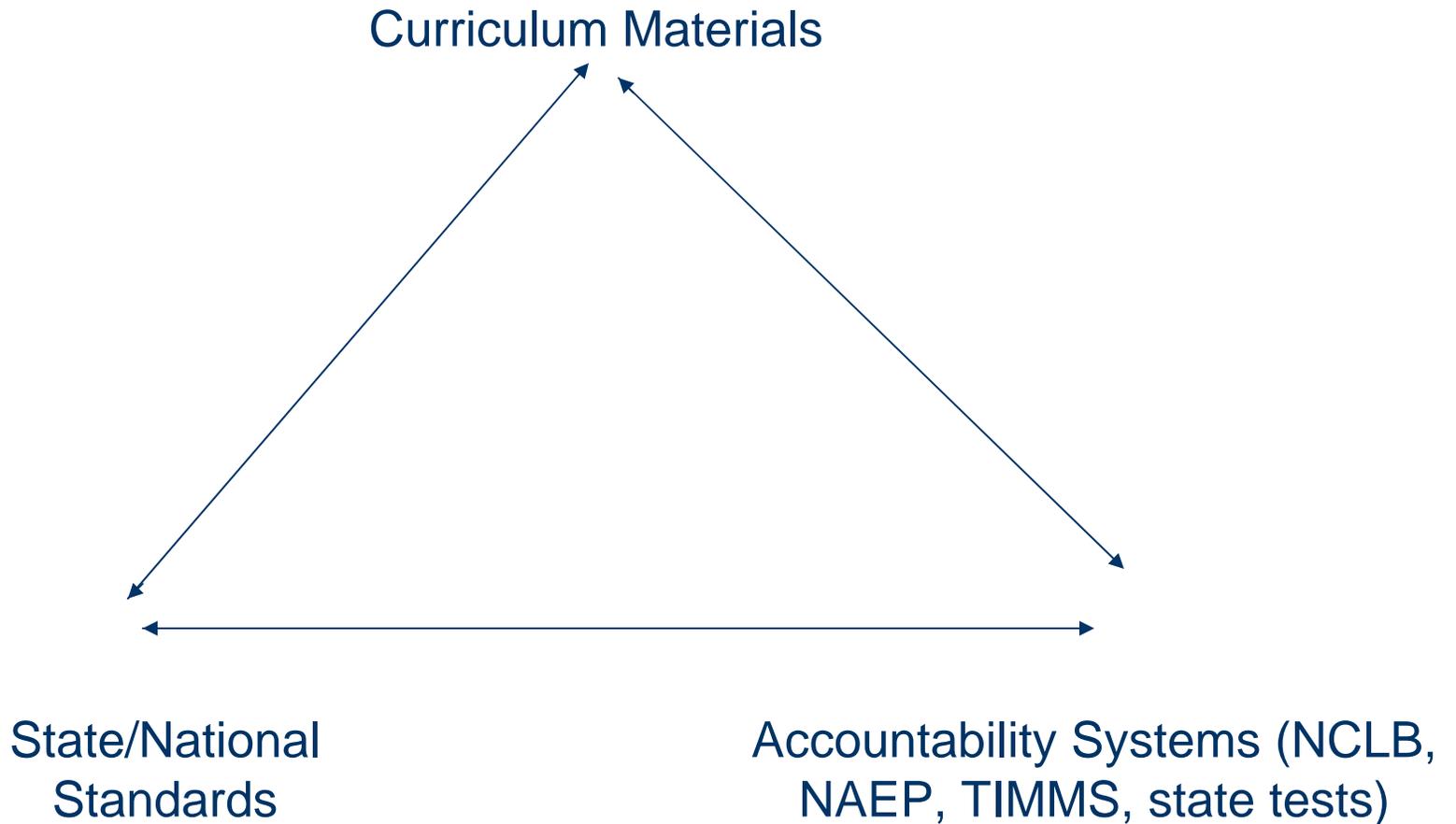
NOTE: Observed differences are not necessarily statistically significant. Detail may not sum to totals because of rounding.
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Science Assessment.

Facts about Science Achievement: No Child Left Behind

- NCLB requires that by 2007–08, states must administer annual science assessments at least once in grades 3–5, grades 6–9, and grades 10–12.
- While schools are not yet accountable for the science tests, the US Dept of Ed (under the American Competitiveness Initiative), is calling for NCLB science assessments to be part of a state accountability systems for adequate yearly progress (AYP).

Source: *NSTA Legislative Update*: February 6, 2006; <http://silverchips.mbhs.edu/inside.php?sid=6505>;
U.S. Department of Education spokeswoman Valerie Smith

Why are curriculum materials important?



NYSED Science Initiative

Mission:

- Create a statewide learning community to support student achievement of the learning standards in science, leading to a scientifically literate population.

Vision:

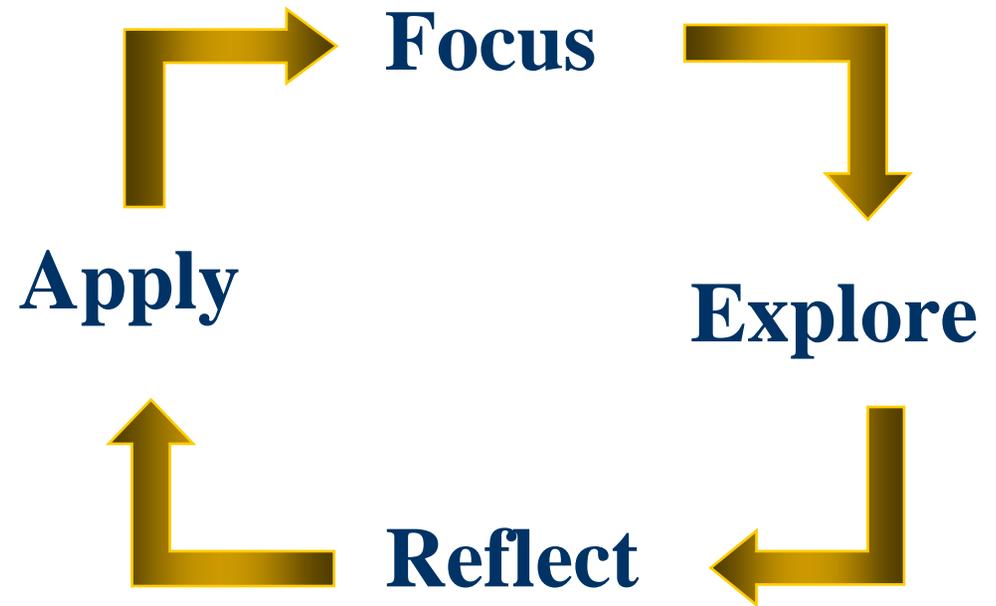
- Ensure the learning of science for all PreK-12 students by providing **E**quitable access to **E**xemplary teachers, **E**ffective inquiry-centered curriculum and instruction, standards-based assessments, and a wealth of resources and community support.

Question

What are the characteristics of an Exemplary, Equitable, and Effective science program?

Exemplary science instruction should...

- Help students build on their prior knowledge.
- Engage students in scientific inquiry.
- Provide opportunities for reflection and problem solving.
- Help students to actively construct knowledge and transfer it to new situations.
- Based on research about how students learn.



Source: www.nsrconline.org

...attend to students' learning progressions

- Incorporate a progression of learning activities (NSF, 2005) that develop modeling, engineering design, and inquiry over time and in the context of national life, earth, physical science, and engineering / technology standards (AAAS 1993; NRC 1996, ITEA, 2000) and the Science Framework for the 2009 National Assessment of Educational Progress (WestEd, Council of Chief State Officers, 2006; retrieved: <http://www.nagb.org/pubs/pubs.html>).

...develop students' scientific reasoning skills in the context of these standards (NSRC)

Grade Level	Observing Measuring Identifying Properties	Seeking evidence, Recognizing patterns	Monitoring cause and effect, Extending senses	Designing Experiments
1	✓			
2-3	✓	✓		
4-5	✓	✓	✓	
6-8	✓	✓	✓	✓

...move students along the continuum of inquiry from directed to guided to full inquiry

		<i>Lecture/Reading Video/Demo</i>	<i>Cookbook Lab</i>	<i>Structured Lab</i>	<i>Challenge Lab</i>	<i>Student-Directed Inquiry</i>	<i>Student-Designed Inquiry</i>
Leading the Inquiry Process	<i>Science Concept</i>	TEACHER				Teacher Student	
	<i>Testable Question</i>						
	<i>Materials</i>	TEACHER			Teacher Student		
	<i>Procedures / Design</i>						
	<i>Results / Analysis</i>	TEACHER		Teacher Student		STUDENT	
	<i>Conclusions</i>						

Source: www.seattlescience.com/Inquiry-Continuum-Poster.pdf

(c.f. pg. 29 Inquiry and the National Science Education Standards)

Question

Are inquiry-based science curriculum materials more **E**ffective and **E**quitable than traditional teaching methods?

Answer

Using a variety of research studies we'll try to "make the case" that inquiry-based science programs have positive effects on student outcomes.

Types of Research Used to Study the Effectiveness and Equity of Curriculum Materials

- Content analyses
- Pre-experiments
- Quasi- and true experiments

Source: On Evaluating Curricular Effectiveness: Judging the Quality of K-12 Mathematics Evaluations (2004).

Available: <http://www.nap.edu/openbook/0309092426/html/index.html>

Source: Subotnik, R. F., Walberg, H. J. (2006). *The Scientific Basis of Educational Productivity*. Greenwich, Connecticut: Information Age Publishing. Review available: <http://edrev.asu.edu/essays/v9n4/>.

Content Analyses

- Content analyses:
 - Tell us whether a curriculum unit aligns with certain content standards.
 - May also provide information on the type of instructional strategies (pedagogy) used to support the learning goals.
 - *But*, should not determine whether the curriculum, when implemented in classrooms, will achieve better learning outcomes for students.

Source: On Evaluating Curricular Effectiveness: Judging the Quality of K-12 Mathematics Evaluations (2004).

Available: <http://www.nap.edu/openbook/0309092426/html/index.html>

Project 2061: Curriculum Analysis

- I. Provide a sense of lesson and unit purpose
- II. Take account of student ideas and misconceptions
- III. Engage students with relevant phenomena
- IV. Help students to develop and use scientific ideas (includes developing meaningful vocabulary in context)
- V. Promote student thinking about phenomena, experiences, and knowledge
- VI. Assess progress
- VII. Enhance the science learning environment and promote curiosity for ALL students

Source: Kesidou, S., & Roseman, J.E. (2002). How well do middle school science programs measure up? Findings from Project 2061's curriculum review. *Journal of Research in Science Teaching*, 39(6), p. 522-549.

Source: www.project2061.org/tools/textbook/mgsci/crit-used.htm

Project 2061 Curriculum Analysis

Instructional Category	Chemistry That Applies*	ARIES** Motion & Forces	McMillian/McGraw-Hill Science*
III. Engaging Students with Relevant Phenomena			
Providing a variety of phenomena	●	●	X
Providing vivid experiences	●	●	X
IV. Developing and Using Scientific Ideas			
Introducing terms meaningfully	●	◐	X
Representing ideas effectively	◐	◐	X
Demonstrating use of knowledge	◑	X	X
Providing practice	●	X	X

● =Excellent, ◑=Very Good, ◐ = Satisfactory, X=Poor

*Source: www.project2061.org/tools/textbook/mgsci/crit-used.htm

**Available from www.gwu.edu/~scale-up under Reports

Pre-experiments

- **Case study**, a type of pre-experiment, asks:
 - How does a particular curriculum material play out in a particular school district at a particular time?
 - What factors influence implementation of the curriculum material in the classrooms or schools of that district?
 - How was the evidence collected?
 - What events were left out and why?
- Other pre-experiments include: **one-group, pre-post design**; and, **comparison of two or more groups, post-test only**.
- By itself, pre-experiments cannot prove that a unit will be effective in any other school or classroom.

Source: On Evaluating Curricular Effectiveness: Judging the Quality of K-12 Mathematics Evaluations (2004).

Available: <http://www.nap.edu/openbook/0309092426/html/index.html>

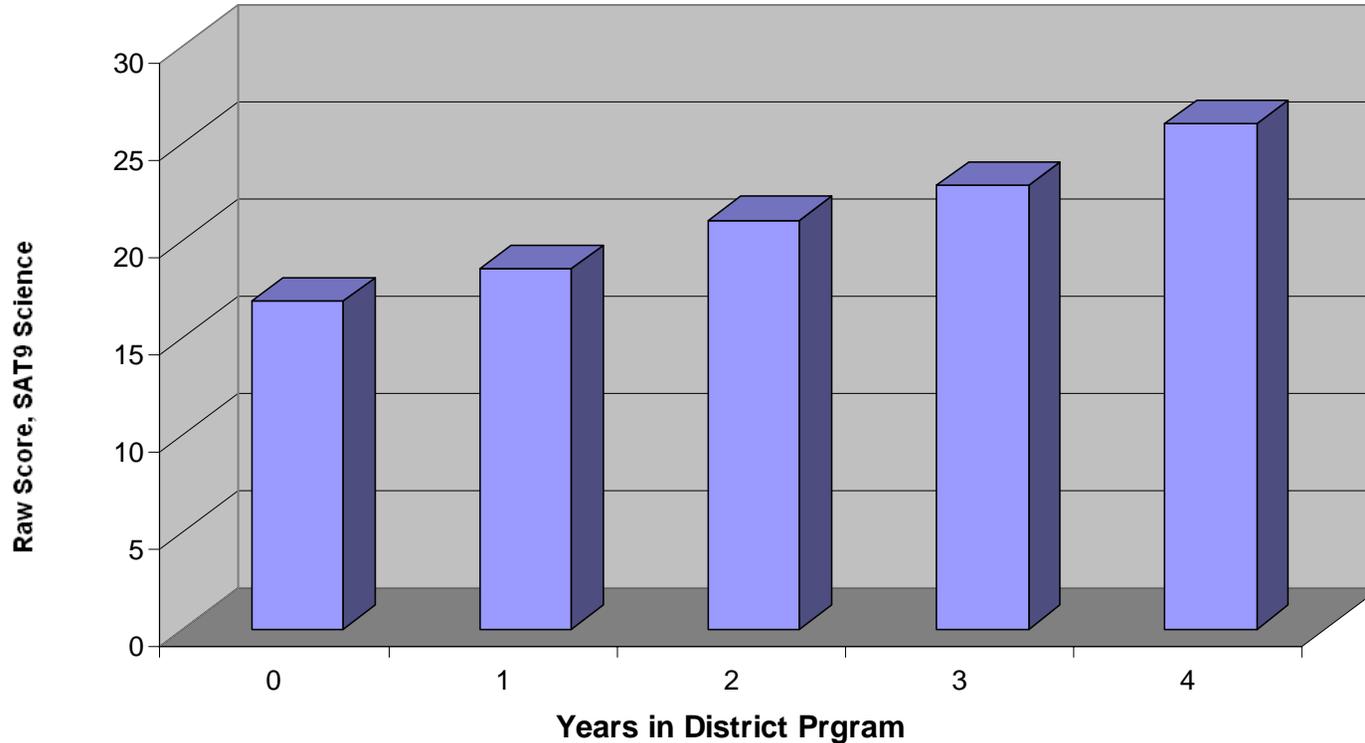
Source: Subotnik, R. F., Walberg, H. J. (2006). *The Scientific Basis of Educational Productivity*. Greenwich, Connecticut: Information Age Publishing. Review available: <http://edrev.asu.edu/essays/v9n4/>.

San Diego State University Valle Imperial Project: El Centro, CA

- The outcomes from the Stanford 9 test conclude that student scores in science, math, reading, and writing are significantly higher for students in inquiry-centered science programs than for students using standard textbooks.
- Students scores in science increased with experience using a combination of FOSS, STC, and Insights materials.

Source: Amaral, O. M., Garrison, L., & Klentschy, M. (2002). Helping English Learners Increase Achievement through Inquiry-Based Science Instruction. *Bilingual Research Journal*, 26, 213-239.

San Diego State University Valle Imperial Project: El Centro, CA



Source: Amaral, O. M., Garrison, L., & Klentschy, M. (2002). Helping English Learners Increase Achievement through Inquiry-Based Science Instruction. *Bilingual Research Journal*, 26, 213-239.

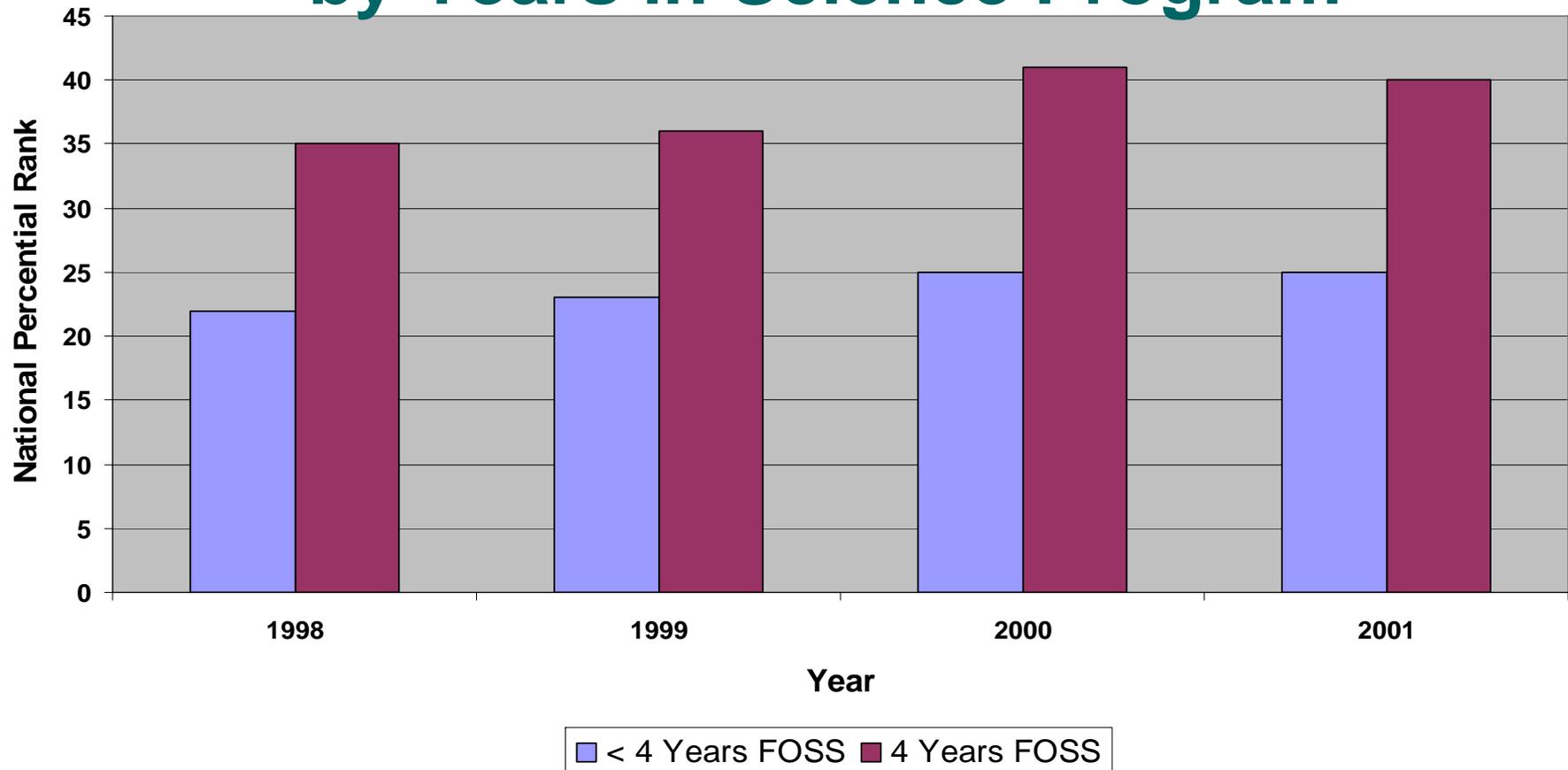
Inquiry-based Science and SAT-9 Reading Comprehension

Inquiry-based science programs, such as FOSS, promote the use of critical thinking skills shown to be associated with increased reading comprehension.

Source: Valadez, J. & Freve, Y. (2002). A Preliminary Summary of Findings from a Study of the Effects of Hands-On/Inquiry-Based Instruction on SAT9 Reading Scores.

Available: sustainability2002.terc.edu/invoke.cfm/page/143

Fresno Unified School District SAT9 Reading Comprehension Scores by Years in Science Program

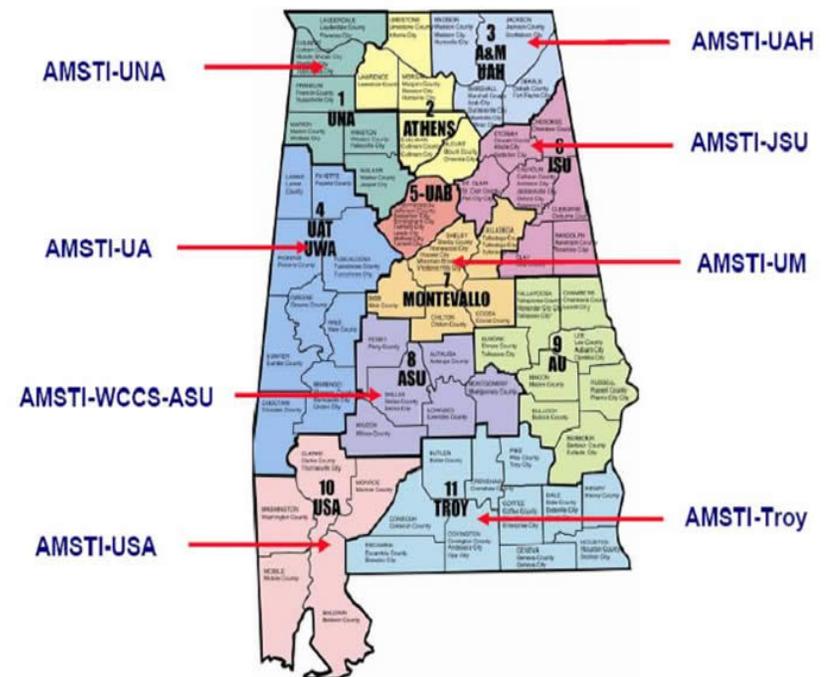


Number of years in an inquiry-based science program is associated with better reading comprehension.

Source: Valadez, J. & Freve, Y. (2002). A Preliminary Summary of Findings from a Study of the Effects of Hands-On/Inquiry-Based Instruction on SAT9 Reading Scores. Available: sustainability2002.terc.edu/invoke.cfm/page/143

Alabama Math, Science, and Technology Initiative (AMSTI)

- Statewide K-12 STEM education reform initiative
- STC, FOSS, Insights, SEPUP used for science, K-8
- As of 2005, 75 AMSTI schools; 265 non-AMSTI schools
- AMSTI teachers receive intensive ongoing PD, and 4 science kits per year

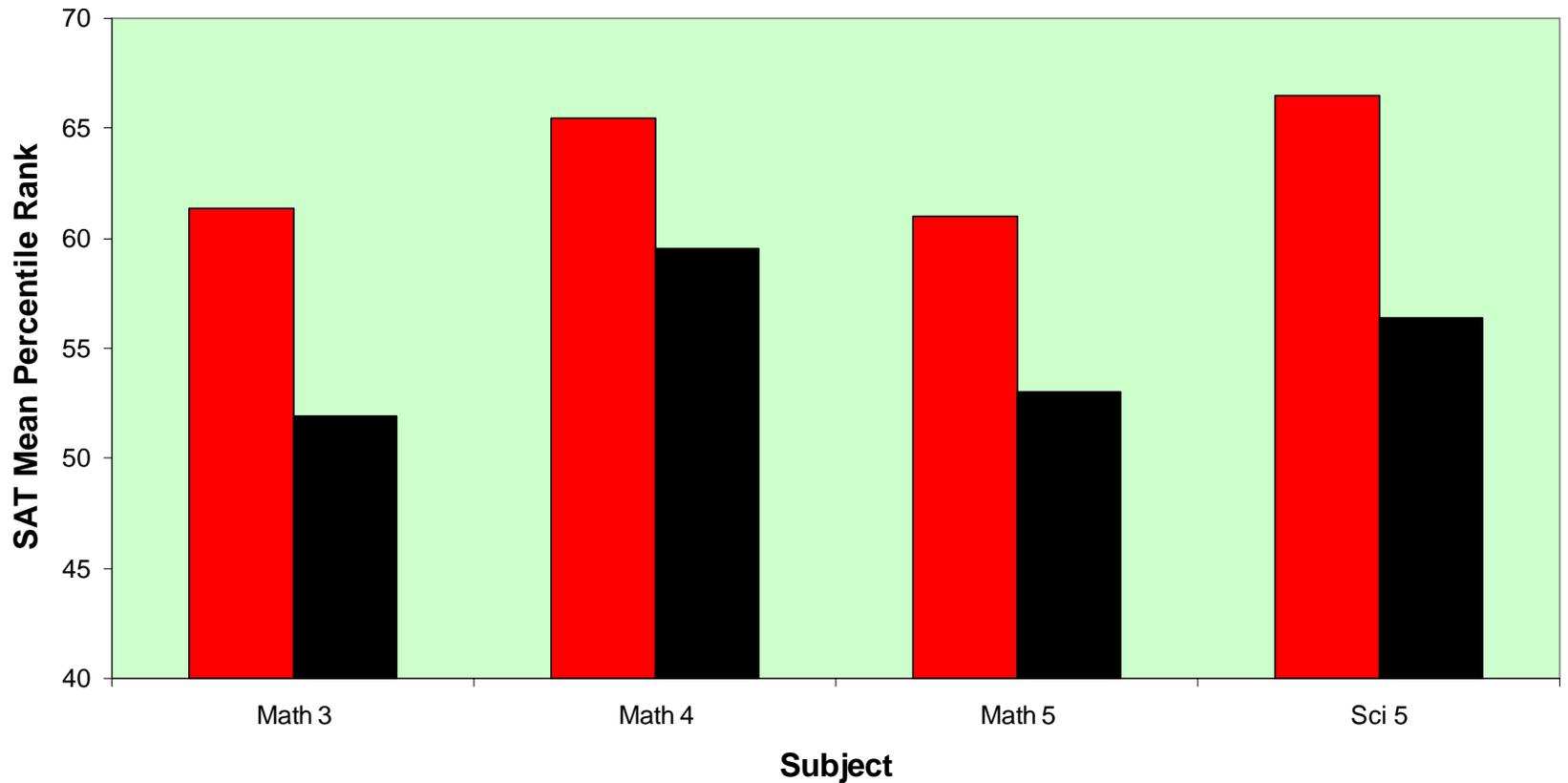


Source: 2005 AMSTI Quality Report, *Impact of AMSTI on Student Performance on Standardized Tests*, Institute for Communication and Information Research at the University of Alabama. Available at: www.amsti.org.

AMSTI vs Matched Comparison

3rd Year of Implementation

(Stanford Achievement Test Version 10)



Red = AMSTI
Schools

Black = demographically
matched control schools

L. A. Unified School District: Integrated Science Study

- High school students using programs such as SEPUP showed gains in student achievement as a result of a two-year integrated science sequence (ISC).
- Disaggregated data show more students enter advanced science classes through the ISC pathway than through the traditional advanced physical science and chemistry pathway.

FIGURE 2.

Percentage of students by ethnicity enrolling in a third year of science after earning a C or better in ICS 2, APS, and chemistry.

Ethnicity	ICS 2	Advanced physical science + chemistry
Hispanic	53.5	35.6
African American	49.4	33.3
Asian	68.8	60.3
Caucasian	50.9	43.4

Source: NSRC © 2000

Source: Scott, G. (2000). Integrated science study. *Science Teacher*, 67 (6), 56–59.
See also: <http://sepuplhs.org/resources/research/>

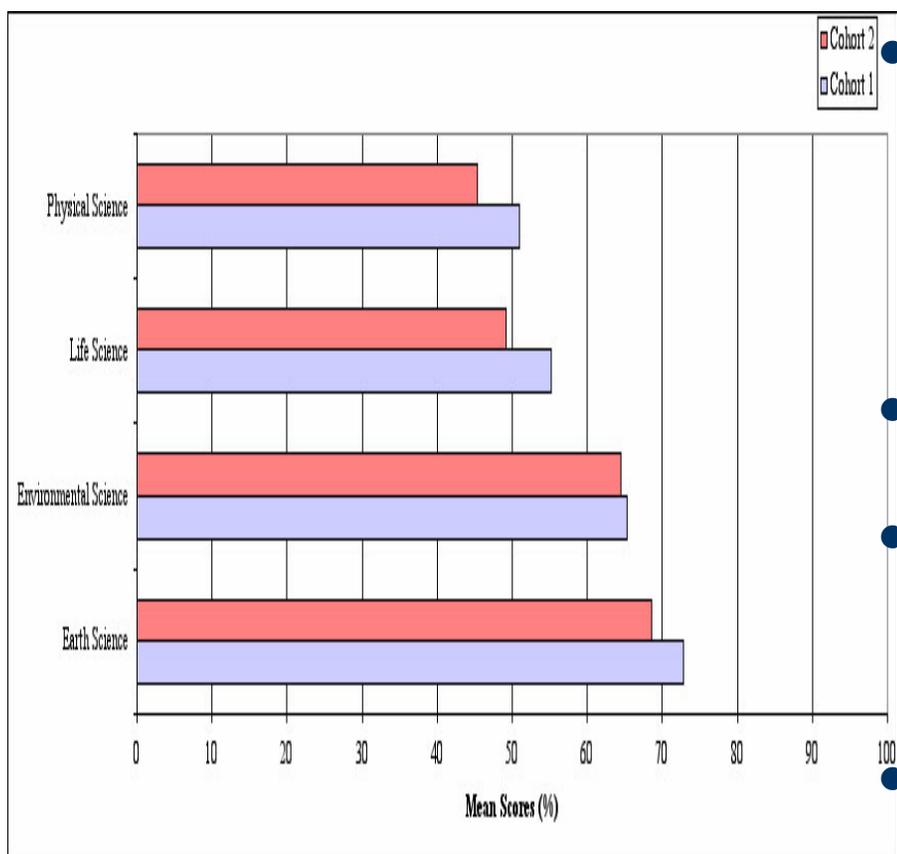
Quasi- and True Experiments

- Compares a given set of schools (or students/classrooms) using the program (Treatment) with same number of schools (or students/classrooms) not using the program (Comparison). NRC recommends at least 5 school pairs. T and C can be the same indiv/group compared over time.
- Groups share similar demographics and previous performance. Called a “true experiment” if random assignment (at the unit of analysis) is used.
- Statistical tests are used to show that if groups are identical at onset, any differences at end resulted from program.

Source: On Evaluating Curricular Effectiveness: Judging the Quality of K-12 Mathematics Evaluations (2004).

Available: <http://www.nap.edu/openbook/0309092426/html/index.html>

ASSET, Inc. University of Pittsburgh



- Univ of Pgh study compared performance of students from two demographically matched ASSET cohorts – districts with 4 to 5 years of involvement vs 1 to 2 years.

- ASSET students use inquiry-based science units exclusively.

- Equivalent groups at onset - No statistical differences on PA System of School Assessment on Reading and Math averages.

- “Long-term involvement in the ASSET project has a wide-ranging impact on learning.”

Science It's Elementary Pennsylvania Initiative

- The Univ of Pgh study in 2001 also showed that ASSET fifth grade students' scores on the TIMSS test were significantly higher than both international and national scores ($p < 0.05$) at both the 4th and 7th grade levels.
- Based on this data, “Science: It’s Elementary”, a “PA Competes” \$10 million/year initiative, is introducing ASSET to 78 schools (1,490 teachers) across 36 counties.

Source: *School Science and Mathematics Journal*, volume 101, Number 8, Dec. 2001(pp. 417-426).

Available: lsc-net.terc.edu/do.cfm/paper/7118/show/use_set-lsc_init

University of Rhode Island

- Compared science achievement of 226 fifth graders from RI LSC districts that have kit-based inquiry curriculum units supported by intensive PD, with data from 173 fifth graders from non-LSC districts using non-kit science materials that have no systematic PD.



Source: Carol O'Donnell © 2000

Young, B. J., & Lee, S. K. (2005). The effects of a kit-based science curriculum and intensive science professional development on elementary student science achievement. *Journal of Science Education and Technology*, 14, 5/6, 471-481.

University of Rhode Island



Source: Carol O'Donnell © 2000

- Administered both pre- and post-tests; demographically matched groups to make certain they were equivalent.
- Kit-based classrooms scored significantly higher than students with non-kit classrooms (even though there were more minutes of science instruction in the non-kit classrooms).

Young, B. J., & Lee, S. K. (2005). The effects of a kit-based science curriculum and intensive science professional development on elementary student science achievement. *Journal of Science Education and Technology*, 14, 5/6, 471-481.

California Institute of Technology

- Studied the degree to which students can “do scientific inquiry.”
- Sampled 1000 fifth graders—hands-on curricula (STC, FOSS, Insights) vs textbook representing 41 classrooms in 9 districts.
- Used four performance assessments: **Paper Towels**, developed by England’s APU 1980s; **Spring**, TIMMS, 1995; **Ice Cubes**, HS item Hopkeiwicz, 1999; and **Flatworms**, Pine et al 2005 (drawn from current literature on inquiry).
- Controlled for cognitive ability, SES, PD hours, and teacher experience. Results were positive for **Flatworms** in favor of hands-on students. No curricular differences found for three physics items.

Source: Pine et al. (2006, May). Fifth graders science inquiry abilities: A comparative study in hands-on and textbook curricula. *Journal of Research in Science Teaching*, 43(5), 467-484.

Johns Hopkins University: Talent Development (TD) Program

- Philadelphia city public schools, as part of Urban Systemic Initiative.
- TD program = integrated curriculum materials (STC/MS, FOSS), materials support, and professional development.
- 3 schools matched on SES and ethnicity as the control.
- Looked for change in SAT9 score and achievement level from 4th – 7th grade.
- TD schools gained 3.5 points (equivalent to 2 stanines) more than students in control schools for every year of exposure to TD.
- Greatest gains were for schools with the highest level of implementation (e.g. support resources—materials, funds).

Ruby, A. (2006). Improving science achievement at high-poverty urban middle schools. *Science Education*, 90, 1005-1027.

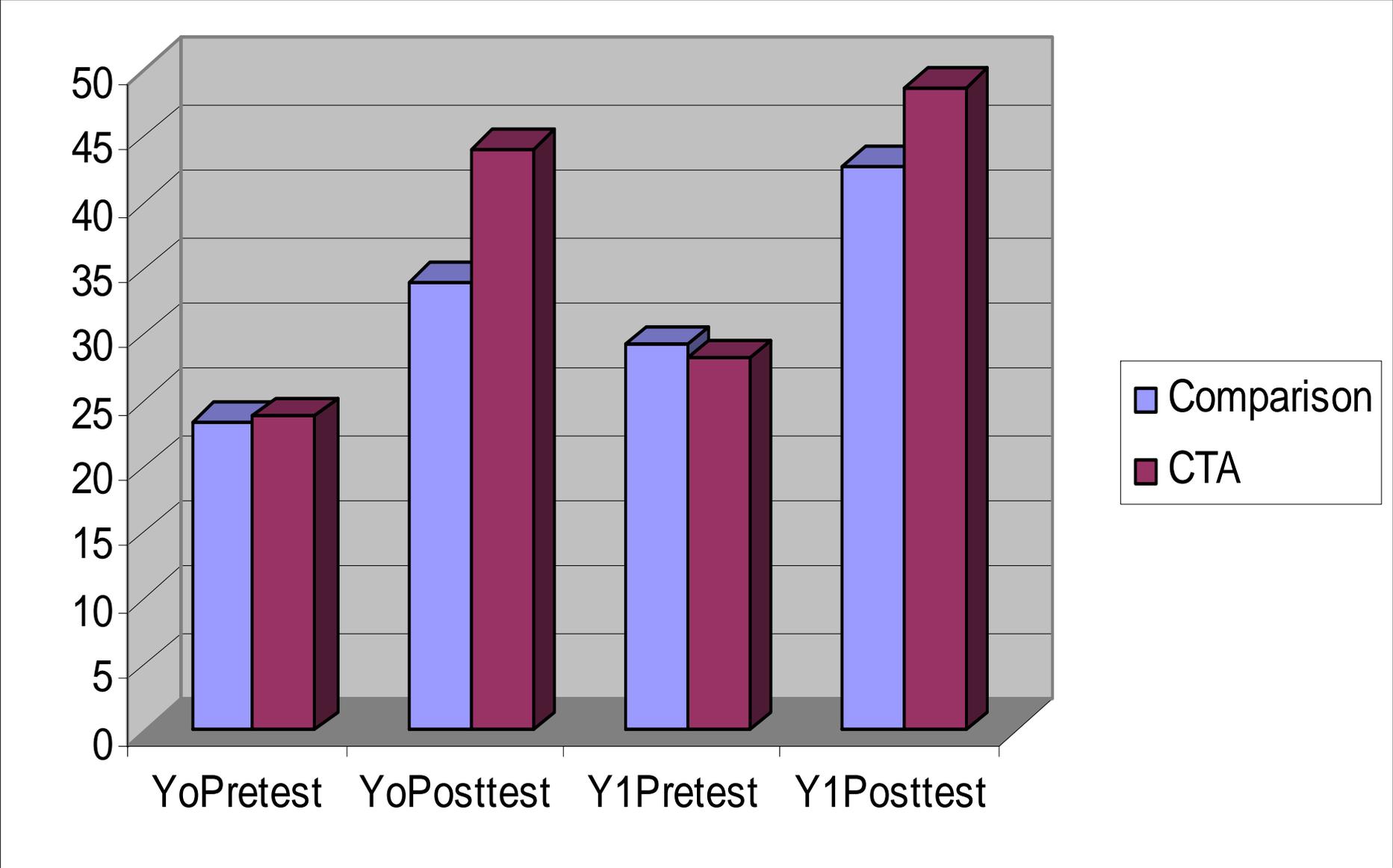
The George Washington University: SCALE-uP Study

- Compared 5 schools **randomly** assigned to *Chemistry That Applies (CTA)* with 5 matched schools not using it.
- Asked: “Do highly rated inquiry-based science units improve student outcomes more than the standard curriculum? Does disaggregating data reveal important differences hidden by overall mean scores?”
- Criterion-referenced assessment with constructed response items administered to measure 8th grade students’ understanding of conservation of matter before and after the 8-week unit. Same tests administered to Comparison group.

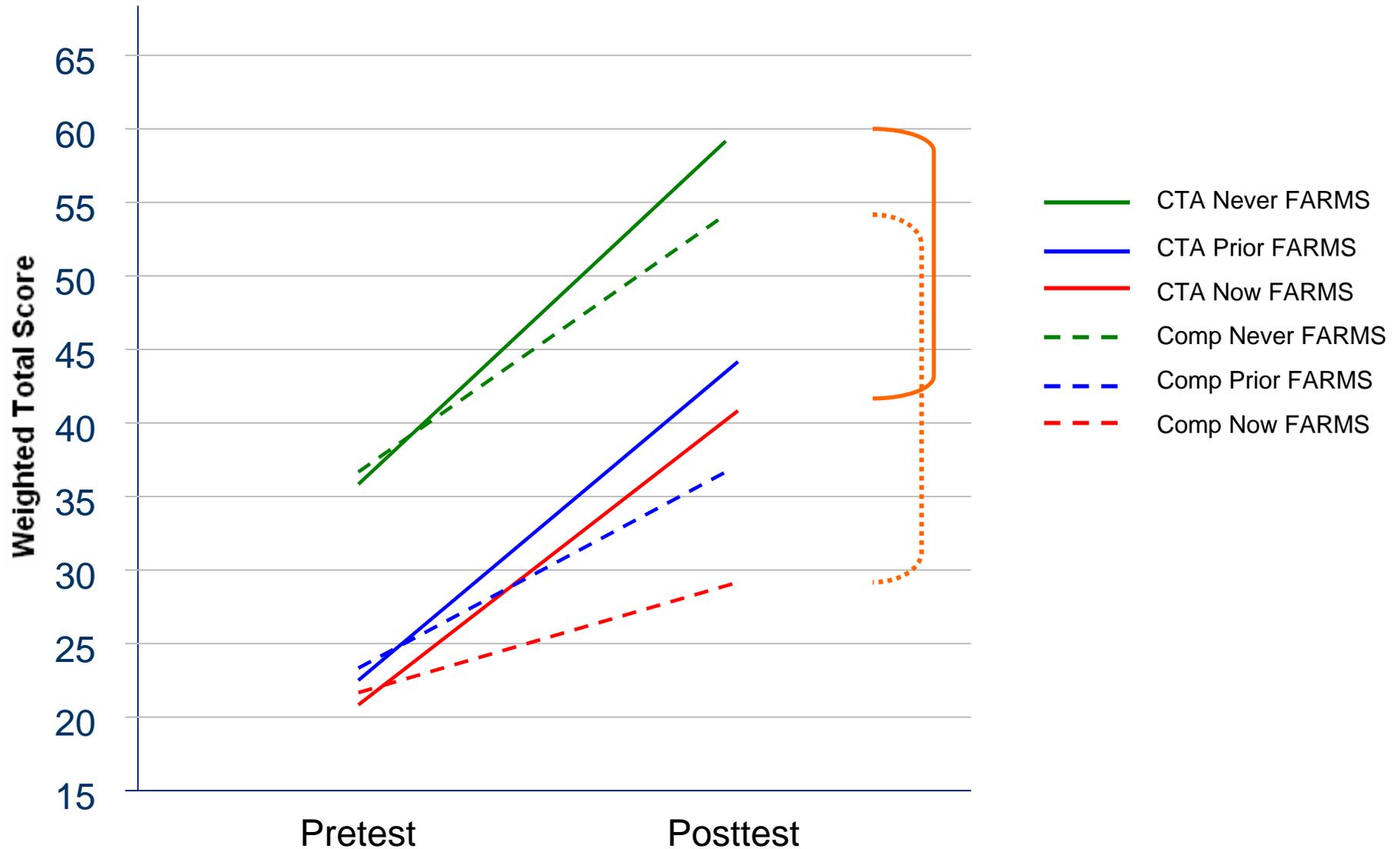
Sources: Lynch, S., Kuipers, J., Pyke, C., & Szesze, M. (2005). Examining the effects of a highly rated science curriculum unit on diverse populations: Results from a planning grant. *Journal of Research in Science Teaching*, 42(8), 912-946.

Lynch, S., Taymans, J. Watson, W. Pyke, C., & Szesze, M. (in press). Scaling up highly rated curriculum units for students with disabilities in mainstream classrooms: Initial findings and implications for scale-up. *Exceptional Children*.

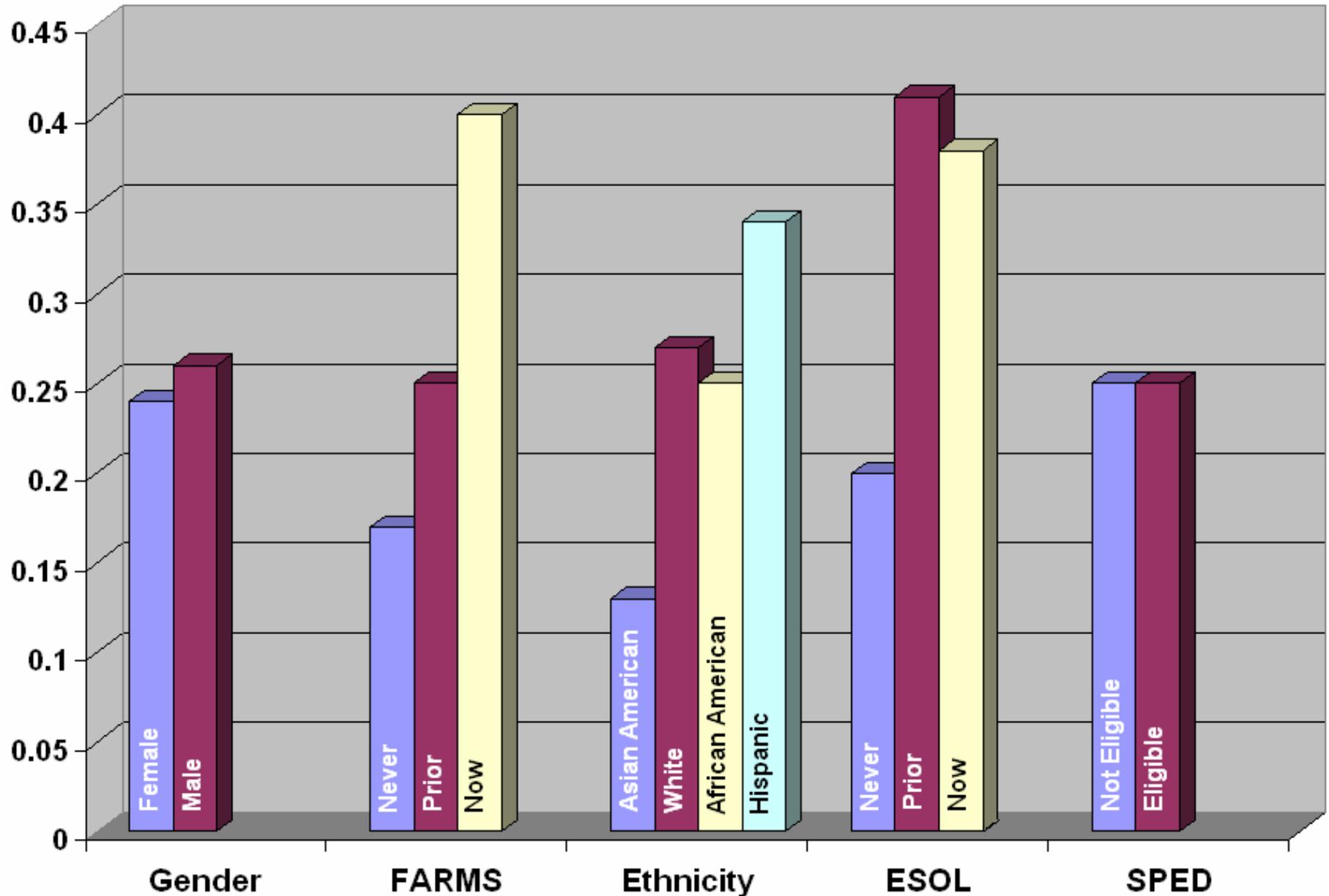
CTA: Mean Scores Year 0 and Year 1



Conservation of Matter Understanding by SES: Year 0 Pre-Post Scores



CTA vs. Comparison: Effect Sizes on Posttest Subgroups for Year 1 (02-03)



Why is disaggregated data important? Equity.

- **NCLB: Adequate Yearly Progress**
 - Achieving gains for *all* subgroups constitutes **AYP** for the state. If even *one* subgroup fails to meet its **AYP** objective, the state fails to meet its **AYP** objective.

Source: Maryland State Department of Education

Why is research about equity and effectiveness important?

- Because NCLB says, “Teachers must use only research-based teaching methods and the schools must reject unproven fads” (NCLB, 2003).
- Programs need to be evaluated in “scientifically based research,” which is defined as “rigorous, systematic, and objective procedures to obtain valid knowledge” (US Dept of Ed, 1998).
- Schools, districts, and states must justify the programs they expect to implement under federal funding (Slavin, 2003).

Source: Slavin, R. E. (2003). A reader's guide to scientifically based research: Learning how to assess the validity of education research is vital for creating effective, sustained reform. *Educational Researcher*, 60(5), 12-16.

Source: www.ed.gov/nclb/methods/science/science.html

Why examine multiple studies?

- "...findings from no single study or even a small set of studies should be taken as the final word on whether a strategy or approach works well. Instead, as many studies as can be found on a given topic should be analyzed. The composite results of those findings should be considered the best estimate of what is known about that topic."

Source: Marzano, R., Pickering, D., & Pollock, J. (2001). Classroom Instruction That Works: Research-based Strategies for Increasing Student Achievement. Alexandria, VA: Association for Supervision & Curriculum Development.

Conclusions

- A growing body of evidence suggests that inquiry-based curriculum materials are effective and equitable.
- However, there will never be a guarantee that a given program will work in a given school, just as no physician can guarantee that a given treatment will work in every case.
- Therefore, it is important to take time to look at the research evidence with some care before making such an important decision. It will be well worth the effort!

Let's Have a Reality Check

- "Magnet: Something you find crawling all over a dead cat."
- "Dew is formed on leaves when the sun shines down on them and makes them perspire."
- "The body consists of three parts--the brainium, the borax and the abominable cavity. The brainium contains the brain, the borax contains the heart and lungs, and the abominable cavity contains the bowels, of which there are five - a, e, i, o, and u."

Contact Information

Carol O'Donnell
Research Scientist
Director, Scaling up Curriculum for Achievement,
Learning, and Equity Project (SCALE-uP)
The George Washington University
2134 G. St., NW
Washington, DC 20052

Telephone: 202-994-4182
Fax: 202-994-0692
E-mail: codonnel@gwu.edu
Web: www.gwu.edu/~scale-up